Ph.D. Title Proposal

on

Development of a Privacy-Preserving Learning Based Framework for Healthcare

Submitted in partial fulfillment of the requirement
for the award of the Degree of

Doctor of Philosophy

in

Computer Engineering

by

Ms. Pramila Prashant Shinde

under the guidance of

Dr. Vidyullata Devmane

Department of Computer Engineering
Mahavir Education Trust's
Shah and Anchor Kutchhi Engineering College
Chembur, Mumbai - 400 088
University of Mumbai
August - 2025

PROPOSAL FOR TOPIC APPROVAL TO BE SUBMITTED TO THE UNIVERSITY OF MUMBAI FOR THE DEGREE OF **Ph.D. (SCIENCE AND TECHNOLOGY)** IN THE SUBJECT OF **COMPUTER ENGINEERING**

Title of Topic	:	Development of a Privacy-Preserving Learning Based Framework for Healthcare
Name of the Student	:	Ms. Pramila Prashant Shinde
Basic Qualification	:	M.Tech. (Computer Engineering) B.E. (Computer Engineering)
In the Case of Ph.D. student Year of Passing of PET/NET/SET/GATE (Teaching Experience)	:	PET - 2024 (70/100)
Name of the Research Supervisor / Guide	:	Dr. Vidyullata Devmane
Date of Superannuation / Retirement	:	26/01/2037
Name of the Research Center	:	Shah and Anchor Kutchhi Engineering College
Signature of Student	:	
Signature of the Research Supervisor	:	

Date: 04-08-2025

Seal of College

Head / Director / Principal
Research Centre

PRINCIPAL
Shah & Anchor Kutchhi Engineering College
Autonomous Institute
Mahavir Education Trust Chowk,
W. T. Patil Marg, Next to D-Mart,
Chembur - 400088

INDEX

Sr. No.	Content	Page No.
1	Introduction	1
2	Literature Survey	3
3	Research Hypothesis	10
4	Aim & Objectives	10
5	Materials & Methods	11
7	Expected Results/Outcomes	12
8	References	13

Statement by the Candidate

I wish to state that the work embodied in this synopsis titled "Development of a Privacy-Preserving Learning Based Framework for Healthcare" forms my own contribution to the work carried out under the guidance of Dr. Vidyullata Devmane at the Shah and Anchor Kutchhi Engineering College. I declare that this written submission represents my ideas in my own words and where others ideas or words have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission.

(Candidate Signature)

Name: Ms. Pramila Prashant Shinde

Roll Number: 124PHCM1007

Introduction

With the growing reliance on personal data for access and processing, concerns regarding data privacy have become more prominent. These concerns arise because such data often contains sensitive information. Even if this information does not directly identify individuals, advanced AI techniques can infer private details. One widely adopted solution is anonymizing datasets containing sensitive attributes.

A notable example occurred in 2006 when Netflix released an anonymized dataset of 100 million movie ratings to support a recommendation algorithm competition. However, in 2008, researchers managed to de-anonymize the data by cross-referencing it with publicly available datasets, revealing personal user details such as political and religious preferences.

Breaches of privacy can significantly harm individuals—potentially resulting in reputation damage, identity theft, discriminatory outcomes, or legal repercussions.

To mitigate such risks, privacy must be integrated through two major approaches: privacy by design, which embeds protective mechanisms during system development, and privacy by policy, which provides operational guidelines. Furthermore, mechanisms like transparency on data usage and adherence to privacy regulations are essential. Implementing privacy by design within AI systems requires a thorough understanding of legal and technical requirements at every development stage, as privacy risks exist throughout the AI lifecycle.

To ensure protection, privacy-preserving strategies must be adopted from the outset and continually improved by analyzing current solutions and addressing their limitations.

The introduction of powerful large language models (LLMs) like ChatGPT introduces new privacy challenges due to their ability to process and possibly expose personal data. Addressing these challenges requires robust privacy policies aligned with laws such as the GDPR. Routine audits, using datasets devoid of identifiable information, and techniques to avoid unintended data exposure are crucial steps.

The 32nd International Conference of Data Protection and Privacy Commissioners emphasized that privacy should be embedded throughout the entire design, operation, and management cycle of any system. Therefore, developing privacy-conscious AI requires identifying and addressing privacy vulnerabilities at each phase of system development. In addition to technical safeguards, compliance with relevant laws and regulations must be ensured. Frameworks like the GDPR define personal data broadly, covering any information that could potentially identify an individual [21].

Federated Learning and Data Partitioning

Federated Learning (FL) is typically categorized into three types based on how data is partitioned: horizontal FL, vertical FL, and federated transfer learning (FTL).

- In horizontal FL, clients have datasets with the same features but different individual records. For example, two hospitals may track the same medical variables but for different patients. Each institution trains a model locally and shares model updates (not raw data) for global aggregation, commonly using algorithms like FedAvg.

- Vertical FL involves clients that share common user samples but hold different feature sets. For instance, one hospital might have patient data related to heart treatment, while another has records of the same patient's checkups. In this case, entity alignment methods are applied, and encrypted collaborative learning is performed using the overlapping data.

- Federated Transfer Learning (FTL) is similar to vertical FL but designed for scenarios where neither features nor samples are fully shared. A pre-trained model is adapted to new tasks

using disjoint datasets from multiple sources, enabling cross-domain knowledge transfer without requiring full feature or sample alignment. For instance, hospitals with unique patient demographics and test types can still collaboratively develop cancer detection models using transfer learning techniques [30].

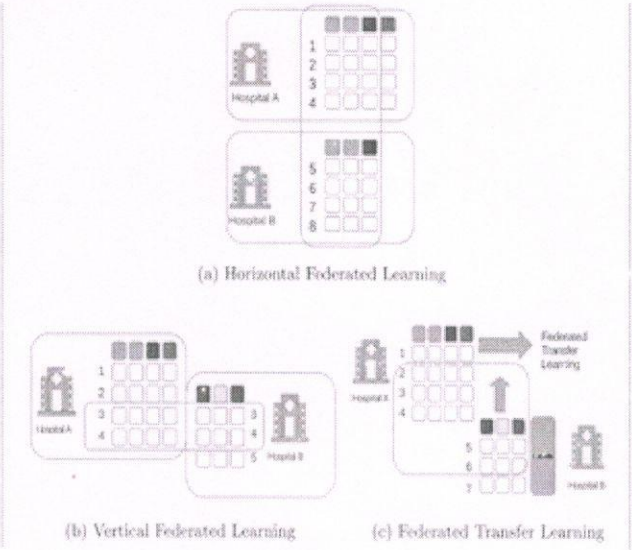

Figure 1: Federated Learning Architecture [30]

Figure 1 illustrates a complete overview of data partition in FL.

Privacy in AI and Healthcare:

Ensuring privacy in AI requires close oversight of the entire data pipeline—not just to secure personal information, but also to reduce bias and enhance model quality. In healthcare, the transition to Electronic Health Records (EHRs) underpins this effort. EHRs are digital documents that store comprehensive patient data, including demographics, medical history, allergies, test results, and treatments. These records are regularly updated by medical professionals and are central to modern healthcare infrastructure [37][38].

Literature Survey

There are more than 60 research papers surveyed by searching with keywords such as privacy, AI, learning in IEEE, Springer, and Sciencedirect databases. There are total 20 research papers selected based on the learning techniques used, or privacy-preserving techniques used, or handling attacks, and applications. Reference No. , Research paper title, Contribution of the author, Datasets used, Techniques used, and Limitations/Gaps are summarised in the table below:

No.	Paper Title / Year	Contribution / Datasets Used	Techniques / Algorithms / Attacks	Remarks / Issues / Gaps
[1]	Evaluation of Privacy-Preserving Support Vector Machine (SVM) Learning Using Homomorphic Encryption (2025)	Fully homomorphic encryption (FHE) encrypts data for machine learning models	Homomorphic encryption, SVM-poly, SVM-linear	High computational cost, latency and memory consumption, need for alternative cryptographic schemes
[2]	Robust and Secure Federated Learning with Verifiable Differential Privacy (2025)	DP noise impact on training accuracy and Byzantine-resilience. Datasets: MNIST, FMNIST, CIFAR-10	Label flipping, model poisoning, backdoor injection attacks. Dynamic noise adjustment, hardware-accelerated verification	Guiding parameter selection, enhancing verification
[3]	RoPA: Robust Privacy-Preserving Forward Aggregation for Split Vertical Federated Learning (2025)	Modified SNIP algorithm for verifying embedding vectors. Datasets: MNIST, Fashion-MNIST, CIFAR-10, Tiny-Imagenet, PathMNIST	Embedding poisoning and free-riding attacks. CNN-based methods	Improving VFL robustness, reducing verification costs, handling malicious active parties

[4]	Encrypted Federated Search Using Homomorphic Encryption (2025)	Execute encrypted queries across distributed databases, without the decryption of the data, thus preserving end-to-end confidentiality. Datasets: Crime report dataset	TenSEAL for applying CKKS and BFV encryption schemes	Scalable and secure, private crime data collaborations of law enforcement
[5]	Examining the Robustness of Machine Learning-based Phishing Website Detection: Action-Masked Reinforcement Learning for Automated Red Teaming (2025)	An action-masked reinforcement learning (RL) to realistically emulate evasion attacks and evaluate the robustness of ML-based phishing website detectors	ML-based detectors: Logistic Regression, Random Forest, and Convolutional Neural Networks	Testing the robustness of the detectors
[6]	Analyzing Federated Learning through an Adversarial Lens (2019)	Confidence on malicious objective (poisoning) and the accuracy on validation data (targeted) on the global model, Comparison of weight update distributions for benign and malicious agents. Datasets: Fashion-MNIST, UCI Adult Census dataset	Model poisoning attacks are effective than data poisoning attacks. Fully connected neural network	Vulnerability of the federated learning setting, develop effective defense strategies
[7]	A Novel Edge-based Multi-Layer Hierarchical Architecture for Federated Learning (2021)	Simulations and the final accuracy and loss of the obtained models compared with the traditional FL approach	Multi-layer hierarchical architectures for FL	Model aggregations at different layers, different computation capabilities and communications

[8]	HED-FL: A hierarchical, energy efficient, and dynamic approach for edge Federated Learning (2023)	A novel, energy-efficient, and dynamic FL-based approach considering a hierarchical edge FL architecture called HED-FL, sustainable learning paradigm using model parameters aggregation at different layers and considering adaptive learning rounds at the edge to save energy but still preserving the learning model's accuracy. Datasets: CIFAR-10 dataset	VGG-11 is an 11-layer convolutional network	Aggregation method used affects the quality of the final model in vertical and horizontal data partitioning
[9]	ProVFL: Property Inference Attacks Against Vertical Federated Learning (2025)	Attacks can achieve inferences with low estimation errors as low as 1%. Datasets: Adult, Census, Bankmk, Lawschool, Texas	ProVFL-B and active attacks as ProVFL-M to enhance performance of DC	VFL distribution-level privacy leakage has not been thoroughly explored
[10]	VerifyVFL: Practical Verifiable Vertical Federated Learning (2025)	The first practical verifiable neural network-based VFL scheme. Datasets: Boston House Dataset, MNIST, CIFAR-10. ImageNet dataset under DenseNet121, ResNet50, VGG19. gradient proofs	Support the verification of non-linear gradients	Not suitable for certain loss functions

[11]	Theoretically Unmasking Inference Attacks Against LDP-Protected Clients in Federated Vision Models (2025)	Practical evaluations on federated vision models confirm considerable privacy risks, revealing that the noise required to mitigate these attacks significantly degrades models' utility. Datasets: CIFAR10, CIFAR100 and ImageNet	For ResNet, pretrained foundation models provided by the HuggingFace	Stronger privacy safeguards, robust defenses against AMI attacks, vulnerable to inference attacks in FL with dishonest servers and clients
[12]	AdeptHEQ-FL: Adaptive Homomorphic Encryption for Federated Learning of Hybrid Classical-Quantum Models with Dynamic Layer Sparing (2025)	AdeptHEQ-FL delivers competitive accuracy and efficiency compared to prior QFL approaches. Datasets: CIFAR-10, SVHN, FashionMNIST dataset	Adaptive homomorphic encryption for federated learning of hybrid classical-quantum models	Extend encryption coverage, test on physical devices, more complex datasets
[13]	Federated Learning-Enhanced Blockchain Framework for Privacy-Preserving Intrusion Detection in Industrial IoT (2025)	A smart contract-enabled blockchain system records model updates and anomaly scores to establish accountability. Datasets: ToN-IoT, N-BaIoT	Model poisoning and Byzantine attacks	Through reputation-aware aggregation schemes also remains a promising direction
[14]	Privacy-preserving Mutual Authentication Protocol for Federated Learning in Intelligent Transportation Systems (2025)	Privacy-preserving Decentralized Identifiers (DIDs) and Verifiable credentials (VCs) together with lightweight cryptographic operations	Protection from several attacks	Efficient authentication solution for real-world deployment

[15]	Hierarchical over-the-air federated edge learning (2025)	A cluster-based hierarchical model outperforms the conventional FL. Datasets: MNIST, CIFAR-10 datasets	A cluster-based hierarchical model	Performance is limited by the presence of mobile users (MUs) located far away from the Parameter Server
[16]	Federated collaborative filtering for privacy-preserving personalized recommendation system (2025)	Empirical validation confirms a collaborative filter can be federated without a loss of accuracy hence enhancing the user's privacy. Datasets: MovieLens and an in-house dataset	Privacy-preserving recommendation systems with a federated model	Investigate the effects of attacks and threats
[17]	A decentralized privacy-preserving framework for diabetic retinopathy detection using federated learning and blockchain (2025)	It integrates FL, SMPC, Blockchain, and IPFS for privacy and scalability.	FL-BC-SMPC-SMOTE integrates multiple technologies for DR detection	Limited exploration in healthcare, privacy-preserving techniques require further development and validation
[18]	Secure and Efficient Drug Supply Chain Management System: Leveraging Polymorphic Encryption, Blockchain, and Cloud Storage Integration (2025)	A comprehensive approach to address the challenges of drug supply chain management	A drug supply chain decentralized application (also called SCMapp) within the Ethereum blockchain	Healthcare-drug-application for real-world pharmaceutical supply chain operations

[19]	Federated learning of predictive models from federated Electronic Health Records (2025)	To predict future hospitalizations for patients with heart-related diseases using EHR data spread among various data sources/agents. Datasets: Boston Medical Center	Decentralized framework, the cluster Primal Dual Splitting (cPDS) algorithm	Analysis of cPDS when the graph that connects the agents is time-varying
[20]	MetaFed: Federated Learning among Federations with Cyclic Knowledge Distillation for Personalized Healthcare (2024)	MetaFed treats each federation as a meta distribution and aggregates knowledge of each federation in a cyclic manner. Datasets: OrganAMNIST, OrganCMNIST, and OrganSMNIST	MetaFed with common methods such as FedAvg, to implement a complete federated learning system, including intra- and inter-federations	Extend MetaFed for heterogeneity architectures, apply MetaFed to more realistic healthcare applications

There are review papers which are on various topics such as Federated Learning, Federated Learning applications and Healthcare applications. Limitations and future directions are studied from these papers.

Limitations

1) There is the need for standard privacy metrics and interoperability in AI systems. [21]

2) 3652.1-2020 - IEEE Guide for Architectural Framework and Application of Federated Machine Learning will promote the use of distributed data sources without violating regulations or ethical considerations. [22]

3) There is plenty of scope to work on Dynamic federated learning, Decentralized federated learning, and Scalability of federated learning. [23, 24].

4) Research on more efficient encryption algorithms and to develop the selection strategy of large-scale clients in FL is future direction. There is an important concern to work on new model fusion method in multimodal FL. This will consequently lead to application of FL in various fields. [25]

5) New application areas can benefit from learning algorithms, such as food delivery systems, VR applications, finance, public safety, hazard detection, traffic control, and monitoring, etc. [26]

6) Learning models in more complex scenarios deserves future work, including label concept shift, feature concept shift, label space heterogeneity, and feature & label space heterogeneity. [27]

7) Research paper gives future direction as scalability, security, and regulatory compliance and the challenges associated with communication rounds, customization, absence of labels, resilience. Due to noise which is added has concern point with respect to accuracy of the learning algorithm. Sometimes classical encryption-based methods fail to meet the standard security requirements. There is need to improve the learning process in terms of privacy, security, and overall effectiveness. There is the major focus to guarantee the confidentiality of the data and the integrity of the model. [28]

8)In healthacre, it is required to work on personalized medicine, and decentralized data. There is

need to protect patients rights with optimized results. The regulatory framework, such as the General Data Protection Regulation (GDPR), is groundwork, to inculcate confidence and trust in patients and stakeholders alike. [29]

9)The user's privacy must be protected for the heterogeneity and decentralized environments of learning. It is essential to increase accuracy. [30–32]

10)Robustness to mitigate the attacks and reduce computational costs remain a future issue for the federated learning. [33]

11)It is essential to make federated learning effective in distributed and isolated data environment. [34]

12)In healthcare applications, patients data is not Independent and Identically Distributed but diverse and heterogeneous, to perform real-world classification tasks is the challenge. [35]

13)It is required to train the CNN with an imbalanced distributed dataset in learning based framework. [36]

Research Hypothesis

Hypothesis 1:
Null Hypothesis (H_0): Existing privacy-preserving techniques do not significantly improve data privacy.
Alternate Hypothesis (H_1): Existing privacy-preserving techniques do significantly improve data privacy.

Hypothesis 2:
Null Hypothesis (H_0): Developing privacy-preserving learning-based framework does not lead to any significant improvement.
Alternate Hypothesis (H_1): Developing privacy-preserving learning-based framework do lead to significant improvement.

Hypothesis 3:
Null Hypothesis (H_0): A learning-based framework for multimodal healthcare data does not significantly enhance data integration outcomes compared to unimodal approaches.
Alternate Hypothesis (H_1): A learning-based framework for multimodal healthcare data significantly improves data integration outcomes compared to unimodal approaches.

Hypothesis 4:
Null Hypothesis (H_0): The proposed privacy-enhanced system does not result in a significant improvement in system performance.
Alternate Hypothesis (H_1): The proposed privacy-enhanced system results in a significant improvement in system performance.

Aims & Objectives

To design and develop enhanced privacy-preserving learning based framework in healthcare applications.

- To analyze existing privacy-preserving techniques for data processing.
- To develop a privacy-preserving learning based framework.
- To design and develop learning based framework for multimodal healthcare data.
- To analyze the effectiveness of the proposed privacy enhanced system.

Materials & Methods

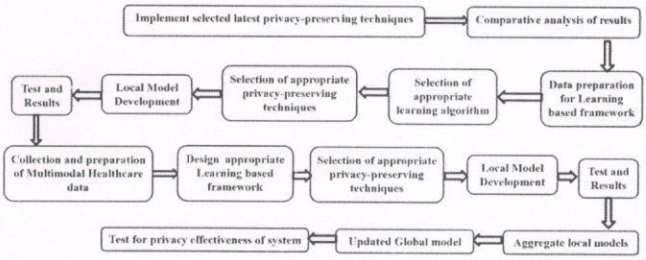
Implement selected latest privacy-preserving techniques
Comparative analysis of results
Data preparation for Learning based framework
Selection of appropriate learning algorithm
Selection of appropriate privacy-preserving techniques
Local Model Development
Test and Results
Collection and preparation of Multimodal Healthcare data
Design appropriate Learning based framework
Selection of appropriate privacy-preserving techniques
Local Model Development
Test and Results
Aggregate local models
Updated Global model
Test for privacy effectiveness of system

Figure 2: Block Diagram

Data Preparation: Appropriate techniques will be applied for data pre-processing and preparation.

Learning based framework: As per data collection, appropriate Learning based framework will be designed.

Multimodal data: Various types of data is required to be collected such as text, images.

Local Model Development: The selection of appropriate learning algorithm for given data is carried out. It has weights specified so as to develop local learning model.

Privacy-preserving Techniques: From literature survey, latest privacy-preserving techniques will be selected. Select and apply privacy-preserving techniques which permit operations on data. It should check for hashing, that is verifiable by third party. It includes encryption, decryption and hash (verification.)

Aggregation of models: Server could be centralized or decentralized and multiple local models will be aggrgated on aggrgation server which is known as Updated global model.

Privacy Effectiveness: The system will be evaluated for its performance and effectiveness of the privacy-preservation.

Test and Results: The results obtained are compared with existing techniques to check for various parameters such as accuracy.

Expected Results/Outcomes

- **Enhanced Privacy:** Secure processing of user data using privacy-preserving technique to protect sensitive information.
- **Collaboration:** Better learning is enabling collaboration across different local models without compromising privacy.
- **User Satisfaction:** Enhance privacy for users through learning which addresses key challenges like accuracy, computational overhead with more robust and generalizable models.

References

[1] W. J. Buchanan and H. Ali, "Evaluation of privacy-preserving support vector machine (SVM) learning using homomorphic encryption," *Cryptography*, vol. 9, no. 2, p. 33, 2025. [Online]. Available: https://doi.org/10.3390/cryptography9020033.

[2] C. Zhang, J. Weng, J. Weng, Y. Zhong, J.-N. Liu, and C. Deng, "Robust and Secure Federated Learning with Verifiable Differential Privacy," IEEE Transactions on Dependable and Secure Computing, pp. 1–18, Jan. 2025, doi: 10.1109/tdsc.2025.3574745. Available: https://doi.org/10.1109/tdsc.2025.3574745.

[3] L. Wang, Z. Zhang, M. Huang, K. Gai, J. Wang, and Y. Shen, "ROPA: Robust Privacy-Preserving Forward Aggregation for Split Vertical Federated Learning," IEEE Transactions on Network and Service Management, p. 1, Jan. 2025, doi: 10.1109/tnsm.2025.3569228. Available: https://doi.org/10.1109/tnsm.2025.3569228.

[4] O. Rathod, A. Baid, and A. K. Cherukuri, "Encrypted federated search using homomorphic encryption," arXiv preprint arXiv:2505.02409, 2025. [Online]. Available: https://doi.org/10.48550/arXiv.2505.02409..

[5] Y. Gao, B. Ampel and S. Samtani, "Examining the Robustness of Machine Learning-based Phishing Website Detection: Action-Masked Reinforcement Learning for Automated Red Teaming," in 2025 IEEE Security and Privacy Workshops (SPW), San Francisco, CA, USA, 2025, pp. 288-293, doi: 10.1109/SPW67851.2025.00041.

[6] A.N. Bhagoji, S.Chakraborty, P.Mittal, and S.Calo, "Analyzing federated learning through an adversarial lens," in Proc. 36th Int. Conf. Mach. Learn. (ICML), PMLR, vol.97, Jun. 2019, pp.634–643.

[7] F. De Rango, A. Guerrieri, P. Raimondo and G. Spezzano, "A Novel Edge-based Multi-Layer Hierarchical Architecture for Federated Learning," 2021 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress (DASC/PiCom/CBDCom/CyberSciTech), AB, Canada, 2021, pp. 221-225, doi: 10.1109/DASC-PICom-CBDCom-CyberSciTech52372.2021.00047.

[8] F. De Rango, A. Guerrieri, P. Raimondo, and G. Spezzano, "HED-FL: A hierarchical, energy efficient, and dynamic approach for edge Federated Learning," Pervasive and Mobile Computing, vol. 92, p. 101804, 2023.

[9] L. Bai, X. Zhang, S. Zhang, Q. Ye and H. Hu, "ProVFL: Property Inference Attacks Against Vertical Federated Learning," in IEEE Transactions on Information Forensics and Security, vol. 20, pp. 6529-6543, 2025, doi: 10.1109/TIFS.2025.3581743.

[10] J. Hou and L. Zhang, "VerifyVFL: Practical Verifiable Vertical Federated Learning," 2025 55th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN), Naples, Italy, 2025, pp. 75-87, doi: 10.1109/DSN64029.2025.00022.

[11] Q. Nguyen, M. N. Vu, T. Nguyen, and M. T. Thai, "Theoretically unmasking inference attacks against LDP-protected clients in federated vision models," in Proc. 42nd Int. Conf. Machine Learning (ICML), Vancouver, Canada, PMLR 267, 2025.

[12] M. A. Jahin, T. R. Fuad, M. F. Mridha, N. Fahad, and M. J. Hossen, "AdeptHEQ-FL: Adaptive homomorphic encryption for federated learning of hybrid classical-quantum models with dynamic layer sparing," in Proc. 1st Int. Workshop on BISCUIT at ICCV, 2025.

[13] A. Ali, M. Husain, and P. Hans, "Federated learning-enhanced blockchain framework for privacy-preserving intrusion detection in industrial IoT," *IEEE Trans. Ind. Informat.*, vol. 21, no. 4, pp. 2451–2462, Apr. 2025.

[14] R. Manoj, S. Joshi, A. S. Aravinthakshan, and C. Tripti, "AdaptPhishSysNet: adaptive phishing detection system for blockchains using machine learning," in Lecture notes in networks and systems, 2025, pp. 389–399. doi: 10.1007/978-981-96-1747-0_32. Available: https://doi.org/10.1007/978-981-96-1747-0_32.

[15] O. Aygun, M. Kazemi, D. Gunduz, and T. M. Duman, "Hierarchical Over-the-Air Federated Edge Learning," ICC 2022 - IEEE International Conference on Communications, May 2022, doi: 10.1109/icc45855.2022.9839230.

[16] M. Ammad-Ud-Din, E. Ivannikova, S. A. Khan, W. Oyomno, Q. Fu, K. E. Tan, and A. Flanagan, "Federated collaborative filtering for privacy-preserving personalized recommendation system," *arXiv preprint arXiv:1901.09888*, 2019.

[17] O. Dib, "A decentralized privacy-preserving framework for diabetic retinopathy detection using federated learning and blockchain," *Results in Engineering*, vol. 26, p. 105456, 2025. [Online]. Available: https://doi.org/10.1016/j.rineng.2025.105456.

[18] M. S. Famous, S. Sayed, R. Mazumder, R. T. Khan, M. S. Kaiser, M. S. Hossain, K. Andersson, and R. Khondoker, "Secure and efficient drug supply chain management system: Leveraging polymorphic encryption, blockchain, and cloud storage integration," *Cyber Security and Applications*, vol. 2025, p. 100103, 2025. [Online]. Available: https://doi.org/10.1016/j.csa.2025.100103.

[19] T.S. Brisimi, R. Chen, T. Mela, A. Olshevsky, I.C. Paschalidis, and W. Shi, "Federated learning of predictive models from federated electronic health records," Int. J. Med. Inform., vol. 112, pp. 59–67, 2018.

[20] Y. Chen, W. Lu, X. Qin, J. Wang and X. Xie, "MetaFed: Federated Learning Among Federations With Cyclic Knowledge Distillation for Personalized Healthcare," in IEEE Transactions on Neural Networks and Learning Systems, vol. 35, no. 11, pp. 16671-16682, Nov. 2024, doi: 10.1109/TNNLS.2023.3297103.

[21] S. Shahriar, S. Allana, S. M. Hazratifard and R. Dara, "A Survey of Privacy Risks and Mitigation Strategies in the Artificial Intelligence Life Cycle," in IEEE Access, vol. 11, pp. 61829-61854, 2023, doi: 10.1109/ACCESS.2023.3287195.

[22] "IEEE Guide for Architectural Framework and Application of Federated Machine Learning," in IEEE Std 3652.1-2020 , vol., no., pp.1-69, 19 March 2021, doi: 10.1109/IEEESTD.2021.9382202.

[23] P. Kairouz et al., "Advances and open problems in federated learning," Foundations and Trends® in Machine Learning, vol. 14, no. 1–2, pp. 1–210, 2021.

[24] B. Liu, N. Lv, Y. Guo, and Y. Li, "Recent advances on federated learning: A systematic survey," Neurocomputing, vol. 597, p. 128019, 2024. [Online]. Available: https://doi.org/10.1016/j.neucom.2024.128019.

[25] J. Wen, Z. Zhang, Y. Lan, et al., "A survey on federated learning: challenges and applications," International Journal of Machine Learning and Cybernetics, vol. 14, pp. 513–535, 2023. [Online]. Available: https://doi.org/10.1007/s13042-022-01647-y.

[26] M. Shaheen, M. S. Farooq, T. Umer, and B.-S. Kim, "Applications of Federated Learning; taxonomy, challenges, and research trends," Electronics, vol. 11, no. 4, p. 670, 2022. [Online]. Available: https://doi.org/10.3390/electronics11040670.

[27] W. Guo, F. Zhuang, X. Zhang et al., "A comprehensive survey of federated transfer learning: challenges, methods and applications," Frontiers of Computer Science, vol. 18, no. 186356, 2024. [Online]. Available: https://doi.org/10.1007/s11704-024-40065-x.

[28] S. Saha, A. Hota, A. K. Chattopadhyay et al., "A multifaceted survey on privacy preservation of federated learning: progress, challenges, and opportunities," Artificial Intelligence Review, vol. 57, no. 184, 2024. [Online]. Available: https://doi.org/10.1007/s10462-024-10766-7

[29] M. S. Amin, S. Ahmad, and W. K. Loh, "Federated learning for Healthcare 5.0: a comprehensive survey, taxonomy, challenges, and solutions," Soft Computing, vol. 29, pp. 673–700, 2025. [Online]. Available: https://doi.org/10.1007/s00500-025-10508-z.

[30] D. N. Sachin, B. Annappa, and S. Ambesange, "Federated learning for digital healthcare: concepts, applications, frameworks, and challenges," Computing, vol. 106, pp. 3113–3150, 2024. [Online]. Available: https://doi.org/10.1007/s00607-024-01317-7.

[31] H. Pan, X. Zhao, L. He et al., "A survey of multimodal federated learning: background, applications, and perspectives," Multimedia Systems, vol. 30, no. 222, 2024. [Online]. Available: https://doi.org/10.1007/s00530-024-01422-9.

[32] Z. Ma, N. I. R. Ruhaiyem, M. Zhang et al., "A review of federated learning technology and its research progress in healthcare applications," Applied Intelligence, vol. 55, no. 765, 2025. [Online]. Available: https://doi.org/10.1007/s10489-025-06627-7.

[33] Y. Feng, Y. Guo, Y. Hou et al., "A survey of security threats in federated learning," Complex & Intelligent Systems, vol. 11, no. 165, 2025. [Online]. Available: https://doi.org/10.1007/s40747-024-01664-0.

[34] N. Saeed, M. Ashour, and M. Mashaly, "Comprehensive review of federated learning challenges: a data preparation viewpoint," Journal of Big Data, vol. 12, no. 153, 2025. [Online]. Available: https://doi.org/10.1186/s40537-025-01195-6.

[35] R. Ballester, J. Cerquides, and L. Artiles, "Quantum federated learning: a comprehensive literature review of foundations, challenges, and future directions," Quantum Machine Intelligence, vol. 7, no. 73, 2025. [Online]. Available: https://doi.org/10.1007/s42484-025-00292-2.

[36] R. K. Chaudhary, R. Kumar, and N. Saxena, "A systematic review on federated learning system: a new paradigm to machine learning," Knowledge and Information Systems, vol. 67, pp. 1811–1914, 2025. [Online]. Available: https://doi.org/10.1007/s10115-024-02257-6.

[37] J. King and C. Meinhardt, "Privacy in an AI era: How do we protect our personal information?" Stanford HAI, Mar. 18,2024.[Online].Available: https://hai.stanford.edu/news/privacy-ai-era-how-do-we-protect-our-personal-information

[38] A. Jain, "Digitalisation of healthcare data through electronic health records will be the next startup boom," The Economic Times, Feb. 21, 2024. [Online]. Available: https://economictimes.indiatimes.com/opinion/et-commentary/digitalisation-of-healthcare-data-through-electronic-health-records-will-be-the-next-startup-boom/articleshow/107891904.cms.